Sensing Scenes With Silicon

DAVID C. MOUNTAIN[1,*] AND ALLYN E. HUBBARD[2]

[1]*Boston University Hearing Research Center and Department of Biomedical Engineering,
44 Cummington St., Boston, Massachusetts 02215; and* [2]*Boston University Hearing Research Center
and Department of Electrical and Systems Engineering and Department of Biomedical Engineering,
8 St. Mary's Street, Boston University, Boston, Massachusetts 02215*

Abstract. Scene analysis, the process of converting sensory information from peripheral receptors into a representation of objects in the external world, is central to our human experience of perception. Through our efforts to design systems for object recognition and for robot navigation, we have come to appreciate that a number of common themes apply across the sensory modalities of vision, audition, and olfaction; and many apply across species ranging from invertebrates to mammals. These themes include the need for adaptation in the periphery and trade-offs between selectivity for frequency or molecular structure with resolution in time or space. In addition, neural mechanisms involving coincidence detection are found in many different subsystems that appear to implement cross-correlation or autocorrelation computations.

Introduction

As we walk in a busy city or even a pristine forest, our senses are bombarded by signals from many sources. The acoustic signals entering our ears are a mixture of sounds produced by many sources as well as innumerable echoes. The photons reaching our retina have been reflected off a complicated montage of clothing, faces, automobiles, and buildings or perhaps off a mixture of leaves, stems, insects, birds, soil, and flowers. Likewise, the molecules reaching our olfactory epithelium may be a mixture of burnt hydro-

carbons, perfume, and the smell of decaying trash or a combination of fragrances from flowers, musk from animals, and byproducts of the breakdown of leaves. We refer to the problem of interpreting this jumble of sensory input and relating it to the physical world as *scene analysis.*

Many of the current ideas about scene analysis in general started with experimental and theoretical work on vertebrate vision. David Marr (1982) introduced a conceptual framework that spanned the entire range of issues from perception down through the physiological mechanisms to the actual underlying computations. The core idea is that sensory systems carry out specific computations that can be described mathematically, and that if these computations are understood, then they can be implemented as computer programs or in electronic hardware.

Our own approach to designing artificial systems for scene analysis follows Marr's lead. We start with physiologically based models that replicate the responses of the sensory receptors and neural structures that appear to be involved with the early stages of sensory processing. These models are then further abstracted to a form in which they can be used as the starting point for the design of very large-scale integrated circuits (VLSI). The VLSI circuits, after fabrication, are then integrated with appropriate sensors, and the outputs are fed to a microprocessor for tasks such as grouping, object localization, and object classification.

Visual Scene Analysis

Visual scene analysis in mammals is believed to take place through a series of parallel pathways (Fig. 1). The image projected by the lens onto the retina is transduced by photoreceptors, and then contrast is enhanced by neural processing before the visual information is split into specialized pathways that appear to extract important features
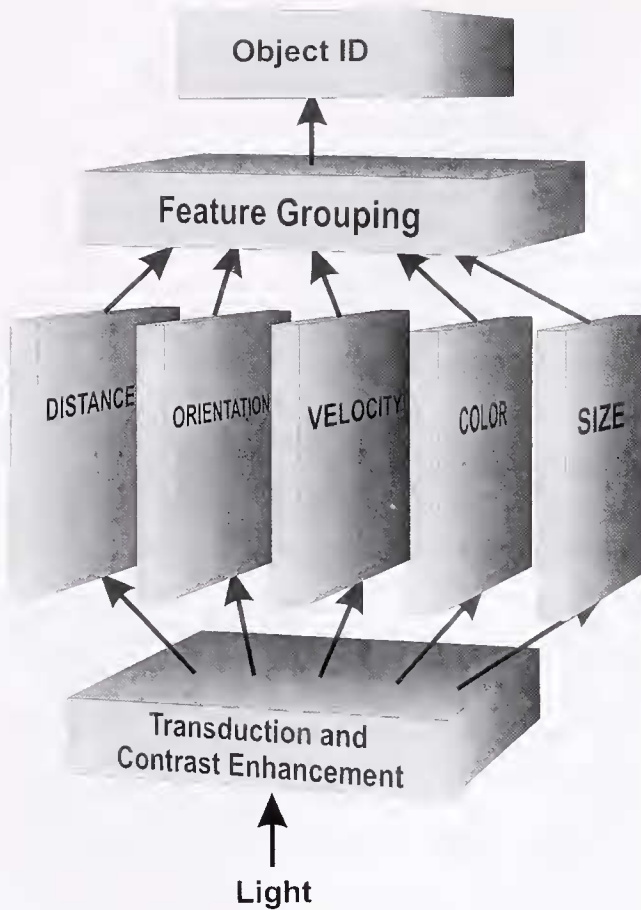
Figure 1. Visual feature analysis consists of first transducing the light projected onto the photoreceptor array and enhancing the contrast of the projected image. This is followed by parallel pathways of feature extraction, the outputs of which are then processed to group related elements to form visual objects.

such as distance, orientation, velocity, color, and size (Marr, 1982). Individual regions of the visual image are analyzed for these different features, and then selected portions are grouped together through selective attention to form visual objects that can be identified.

Similar processes may also be taking place in invertebrates. For example, cells from the third optic ganglion of dragonflies respond selectively to different target classes with properties that are remarkably similar to those of cells from the mammalian visual cortex (O'Carroll, 1993). Also, bees—like mammals—can recognize a familiar shape under a variety of viewing conditions regardless of whether it is initially sensed by color contrast, luminance contrast, or motion contrast (Zhang *et al.*, 1995).

The visual system must be able to cope with the large changes in ambient light level that take place due to time of day, presence or absence of clouds, and moving in and out of the shade. Even with fixed lighting conditions, some parts of the visual scene may be brightly lit while others may be

in the shade. The image projected onto a receptor array is the product of the illumination falling upon the objects within the visual scene, multiplied by the reflectivities of these objects. Since it is the reflectivity (both overall magnitude and spectrum) that provides the useful information about object identity, the visual system needs a method to minimize the effects of varying illumination.

These illumination problems must be dealt with in the first stages of processing, before object formation can take place. The large changes in ambient light level appear to be handled at the receptor level through adaptation. Adaptation is a process whereby the sensitivity of the photoreceptor depends on the time-averaged light level. In biological photoreceptors, biochemical processes provide the needed automatic gain control. The outputs of small groups of photoreceptors are then combined so as to enhance the differences in reflectivity of objects within the scene by using a "center-surround" organization (Fig. 2, column 1). This is done by combining an excitatory input from a receptor or small cluster of receptors with inhibitory inputs from the surrounding neighbors (on-center receptive field) or by combining an inhibitory input from a receptor or cluster with excitatory inputs from the surrounding neighbors (off-center receptive field). Mathematically, the combination of adaptation and center-surround organization is equivalent to performing the combination of local normalization and a two-dimensional second spatial derivative on the output of the receptor array. This process has the effect of emphasizing contrast boundaries in the image. The spatial extent of the receptors contributing to the receptive field can be varied at the design stage to achieve different degrees of resolution (image smoothing). Alternatively, the scene can be processed by parallel pathways each with a different resolution. If appropriate weights are used for the excitation and inhibition, then the center-surround spatial filters can be approximated mathematically as Gabor functions (Weldon and Higgins, 1999). The multi-resolution approach can be thought of as taking a two-dimensional wavelet transform of the image (Porat and Zeevi, 1989).

Distance information is not available to the visual system directly, because the external three-dimensional world is mapped onto a two-dimensional array of receptors. If a three-dimensional internal representation is needed, say for navigational purposes, then the third dimension must be synthesized from the information available from the receptors. If the system has two eyes with overlapping visual fields, then differences due to parallax between the images from the two eyes can be exploited (binocular disparity) to estimate distance; otherwise, vergence or more subtle cues must be used. To estimate binocular disparity, the visual system appears to perform a spatial cross-correlation between corresponding regions of the two retinas (Marr, 1982).

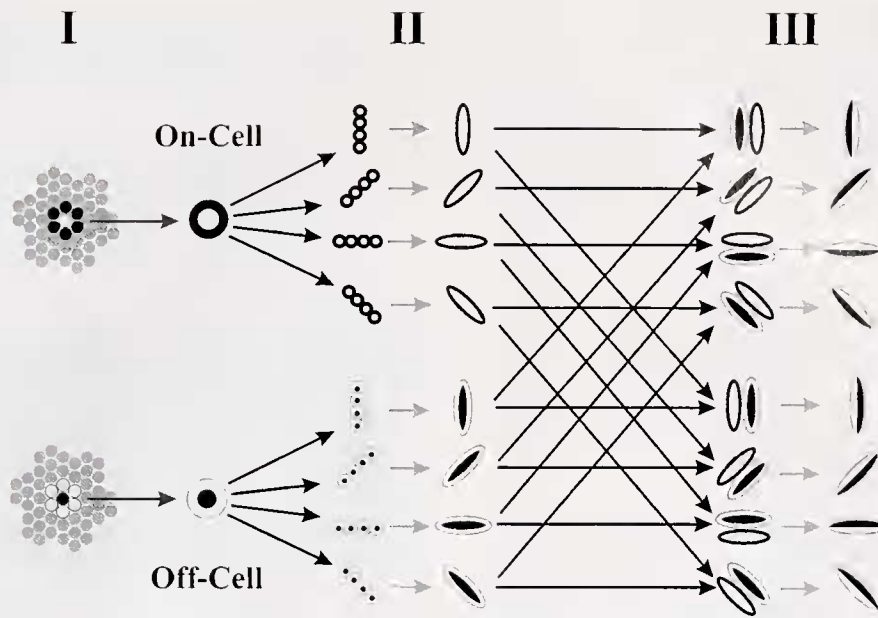Spatial cross-correlation is also used to detect motion.

I   II   III

Figure 2. Orientation processing consists of combining the outputs of cells with center-surround organization (column I) to create oriented receptive fields (column II). These oriented receptive fields are then combined to form oriented edge detectors (column III).

Coincidence detection between the output of a cell and the delayed outputs of other cells with nearby receptive fields is mathematically equivalent to computing the spatial cross-correlation between the current visual frame and a previous visual frame on a region-by-region basis.

Orientation processing involves detecting lines and edges and estimating their angular orientation. Hubel and Wiesel (1962), working with cat visual cortex, showed that detection of oriented edges can be accomplished by a sequence of processing stages that combine the outputs of groups of cells with similar center-surround characteristics. By using groups of cells arranged as short linear arrays, short linear segments of light or dark can be detected (Fig. 2, column II). Different arrays have different orientations (orientation tuning), so that all possible edge segments within a region can be detected. If we then combine the output of pairs of these arrays that are slightly offset from each other and have the same orientation but with one array being of the "on" type and the other being of the "off" type, we have a system that detects edge segments between areas of different reflectivities (Fig. 2, column III). This process can be performed a second time to detect line segments. Higher-level processing can then be used to group the edge or line segments into longer lines and arcs (Pasupathy and Connor, 1999).

We have implemented this type of processing in silicon by designing a set of integrated circuits that implement the processing illustrated in Figure 2 (Hinck and Hubbard, 1999). We do not have space here to go into the details of the silicon implementation, but one significant difference between the biological and silicon system must be mentioned. In biological systems, the information between processing units (cells) is carried by axons that are self routing; in other words, they can work their way through the nervous tissue and find their targets. With silicon processing systems, the wiring problem becomes serious. The processing described within a single column of Figure 2 only requires communication between nearby elements on the chip. However, when we need to move information from one processing level or chip to another (from one column to another in Fig. 2), then we run into problems due to the sheer number of wires involved. To reduce this bottleneck, a technique known as address event representation (AER) is used (Boahen, 2000). When a silicon cell is "excited," it broadcasts its address (identity) to all listeners, which may be a one-to-one or a one-to-many mapping. Each broadcast event is equivalent to the production of a single action potential (spike) in the biological system, and given the bandwidth (speed) of the circuitry we have the ability to transmit the identity of all the spikes from all the cells on a chip. Because the processing is taking place in real time, there is no need to record a time stamp for the events. For simulations that do not run in real time, each event may need both a time stamp and an address.

With AER, signaling takes place only if a spike is generated; this minimizes power consumption because, for a single cell, spikes are relatively rare events. This minimization of power consumption is important, especially for small robots (as well as for biological systems), since low power consumption allows operation for longer periods of time without replenishment of energy stores.

Auditory Scene Analysis

A major challenge in auditory scene analysis is that acoustic signals from different sources can overlap in direction, frequency, and time. We believe that biological systems meet this challenge by dividing up the received signals in frequency and time and—through the use of appropriate grouping principles—enhance the signal-to-noise ratio for individual sources to the point where the bearing and identification of the source can be determined. In many applications, both transient and long-duration signals are of interest. In auditory scene analysis, each frequency band can be analyzed for the presence of specific features, and then the grouping rules can be used to combine information from selected frequency bands to produce the features vector that represents an auditory object.

Audition, unlike vision, has no method by which even two of the three physical dimensions of the external acoustic world can be projected directly onto the receptor array. To determine the direction of a sound source, one either needs to compare signals acquired by directional ears (microphones) with different orientations or compare measurements of pressure taken at different locations in space. In the latter case, the ears or microphones must be spaced sufficiently that the time delay due to the speed of sound is large enough to be sensed or measured. If only two ears or microphones are used, then directional ambiguities are present, but these can generally be resolved through rotation of the head or microphone array. The third dimension (source distance) is much more difficult to estimate in audition. Experiments with human listeners suggest that the ratio of direct to reverberant sound energy may be an important distance cue. How this ratio might be estimated is not clear.

Each frequency channel is analyzed in parallel through the computation of multiple features (Fig. 3). These features are likely to be similar for frequency channels that contain signals from the same sound source and are likely to differ for signals from different sound sources. For example, the differences in time delay between the arrival of the signals (interaural time differences, ITD) as well as differences in intensity (interaural intensity differences, IID) at two sensors will be similar across frequency channels for a single source because these features depend on source direction. Frequency components with similar onsets, offsets, duration, and envelope period are also most likely to be from a single sound source.

For many vertebrates, the head size is sufficient to create significant time delays (ITD) between the ears that can be used for localization; at higher frequencies the head shadow effect is large, producing a significant IID. For very small animals, especially insects, the ears are very close together, making ITD estimation *via* neural circuits impractical, and the animal's size precludes creating a sound shadow. These
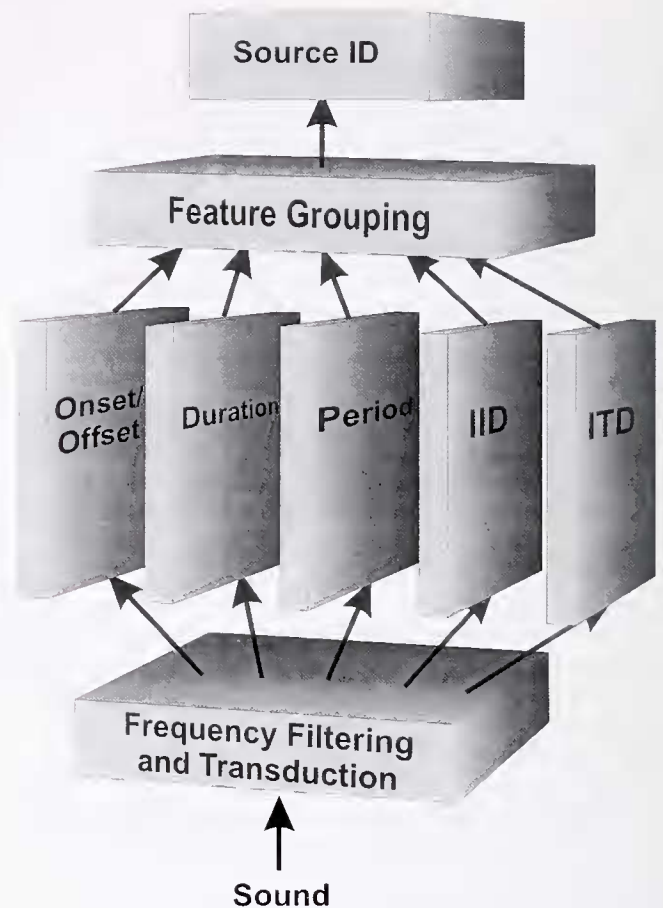

Figure 3.   Auditory feature analysis consists of first filtering and transducing the sound received by the peripheral organs. This is followed by parallel pathways of feature extraction, the outputs of which are then processed to group related elements to form auditory objects.

animals appear to use mechanical or acoustic means, or both, to detect the subtle pressure differences between the two sides of their body (Michelsen, 1998).

As was the case for visual processing, the final step before auditory source identification is the grouping process (Bregman, 1990). In each of the features maps described above, timing information is preserved. This enables the grouping process to use common bearing, as determined by the ITD and IID maps, and synchrony across maps as the major cues for grouping specific components together. This grouping process results in a simplified set of features that includes target direction, the major peaks in the target signal spectrum, and temporal features such as the period of the signal envelope. This set of features can then be compared to stored signatures to complete the identification process. Signatures in this context can be hardwired (acquired through evolution at the species level), learned through experience at the individual level, or derived from a combination of the two methods.

If the system is hardwired, then it is possible to imple-

ment the entire analysis/tracking system with simple circuits. For example, the Webb and Scutt (2000) model of cricket phonotaxis implements pattern recognition and source localization with a system comprising two receptors followed by four neurons. The pattern of interest in this case is the mating call of the male, which is characterized by a limited range of carrier frequencies and a limited range of syllable repetition intervals (SRI) (modulation periods). Filtering for the appropriate carrier frequencies takes place in the hearing organ, and subsequent filtering for SRI takes place using a pair (one for each ear) of output neurons that act as lowpass filters, followed by another pair of neurons that act as a highpass filters. Source localization is accomplished by using directional ears and a combination of excitation and inhibition in the same neurons that perform the highpass filtering.

For auditory scene analysis, it is essential that the filters that perform the frequency separation be designed to have impulse responses that are compact both in frequency and time. The performance measure commonly used to describe this feature is the time-bandwidth product. Simple, single mode resonances, although narrow in frequency, do not have good temporal performance and hence do not have good time-bandwidth products. The impulse response that achieves the theoretical time-bandwidth product limit is a sinusoid with a Gaussian envelope (Gabor function). Such an impulse response is physically unrealizable, but it is possible to combine multiple resonances to create a response that comes close to the ideal. Also, for a general purpose signal processing system, it is generally better to use filters with a constant ratio of bandwidth to center frequency (constant Q) rather than a constant bandwidth like that obtained with a Fourier transform. The widespread use of approximately constant-Q filtering across the ears of many species ranging from bush crickets (Hoy, 1992) to mammals (Javel, 1986) suggests that this approach offers significant survival value. The use of a constant-Q filter bank is very similar mathematically to taking a wavelet transform of the acoustic time signal. It should be noted that most of the acoustic frequencies of biological significance are higher than what most cells can follow, so the filtering is generally done mechanically before detection by the receptor cells. The number of frequency channels may vary from very few in insects (Michelsen, 1992) to hundreds in many vertebrates (Echteler et al., 1994).

Typically this filtering process is implemented in silicon using a cascade of second-order filters with progressively lower resonant frequencies. This cascade is intended to simulate the traveling wave of the mammalian cochlea, which starts in the basal (high-frequency) end of the cochlea and propagates towards the apical (low-frequency) end. For this purpose, subthreshold circuits have been most commonly used (Mead, 1989; Fragniere et al., 1997; Sarpeshkar et al., 1998).

Like the visual system, the auditory system must also deal with a wide range of signal levels. Here again, adaptation (automatic gain control) plays an important role. In mammalian auditory systems the adaptation is specific to each frequency channel (Javel, 1986). In insects, responses of neurons in the central nervous system can also exhibit adaptation (e.g., see Lewis, 1992).

Unlike the visual system, however, the auditory system is processing a very rapidly changing signal, one that often changes much faster than the biological hardware can follow. To circumvent the problem of following high-frequency signals, the receptor cells (hair cells) act as soft half-wave rectifiers (Mountain and Hubbard, 1996) so that at high frequencies they respond to the envelope of the acoustic signal rather than to the fine structure of the signal.

In the auditory system, temporal cross-correlation and autocorrelation-like processing is believed to play an important role (Colburn, 1996; Lyon and Shamma, 1996). In vertebrates, the time delay between the two ears (IID) is an important cue for localization. The combination of neural delay lines and coincidence detection is used to cross-correlate the signals from the two ears for each frequency channel. Periodicity analysis is believed to take place also using delays and coincidence detection. Periodicity analysis no doubt plays an important role for many species from insects to man, because so many communication sounds involve periodic amplitude modulation (AM). Figure 4 illustrates time waveforms in which AM is a prominent feature for a cricket call (panel A) and for a human vowel (panel C). Panels B and D show the results of spectral analysis using a constant-Q filter bank, and except for center frequency and modulation rate, the AM signals are remarkably similar.

Olfactory Scene Analysis

By analogy to the visual and auditory systems, we refer to the problem of identifying and localizing odor sources in complex environments as *olfactory scene analysis*. Unlike vision and hearing, in which the signal propagates *via* wave phenomena, olfaction is characterized by mass transport by currents in water or air and the associated turbulence found in these media (Grasso, 2001). No direct information about source location is present in the received signal, but approximate direction can be estimated by sensing wind or water-flow direction. The only way a source can be located with any certainty is to trace the odor plume back to its source.

In general, individual odor sources release mixtures of compounds into the environment, and the signal at the sensory organ is the result of the mixing of turbulent plumes from multiple sources. Due to the nature of turbulent transport, the plume produced by a single odor source is made up of a series of patches or filaments distributed within the plume; these move past the olfactory organ, creating a series
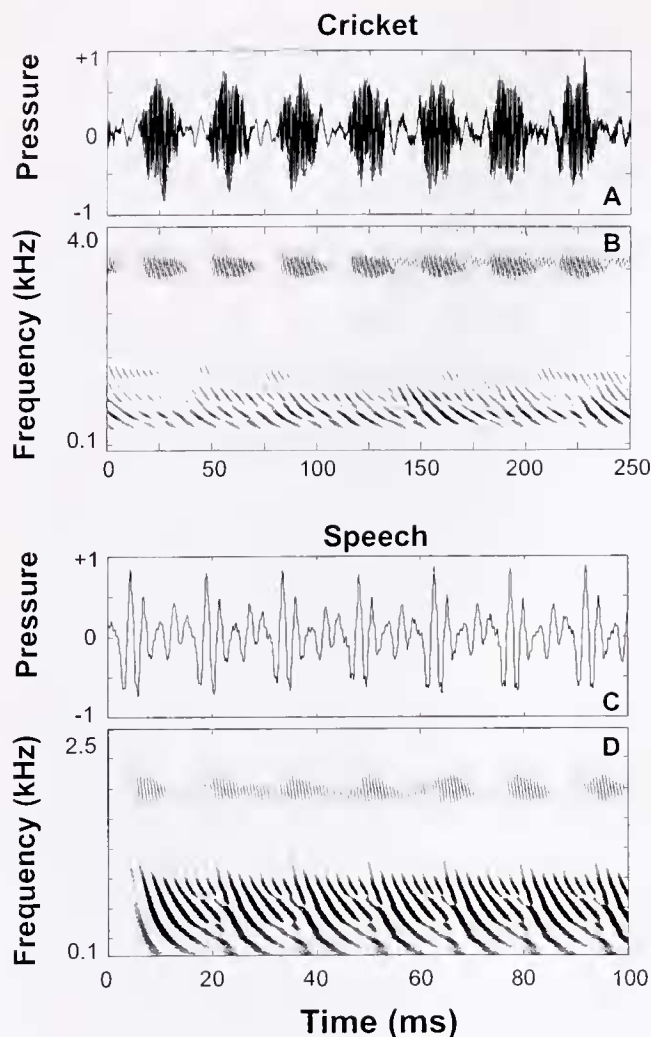
Cricket

centration) level. Olfactory receptors, like their counterparts in the other sensory systems, also exhibit adaptation that adjusts the sensitivity of individual receptors on the basis of background concentration levels. Olfactory systems have many different receptor types, ranging from a few dozen in insects to approximately 1000 receptor types in mammals. Some receptors, mainly those that have evolved to detect pheromones, are extremely selective, but most will respond to a number of different compounds. The higher the selectivity of a receptor, the higher the affinity for the odor molecule and the slower the release of the odor molecule after it has bound to the receptor (Lauffenburger and Linderman, 1993). The relationship between high affinity and slow release comes about because affinity depends on the ratio of the binding to unbinding rates. The rate of binding is limited by the rate at which the odorant can access the binding site, a rate that is similar for all receptors. Affinity, therefore, varies from receptor to receptor, largely due to differences in the unbinding rate. This relationship means
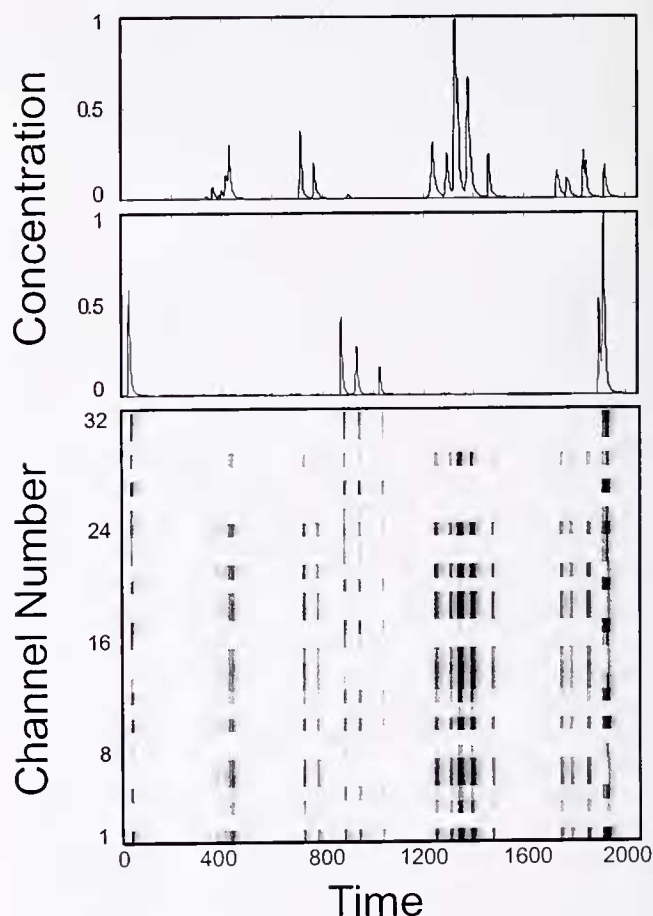
Figure 4. Spectral analysis of animal communication sounds. Time waveforms for a cricket call (panel A) and for a human vowel (panel C) are plotted along with the results of spectral analysis using a constant-Q filter bank (panels B and D).

of odor pulses at the receptors with random arrival times, durations, and amplitudes (Moore and Atema, 1991). The patchy nature of odor concentration signals can be seen in the two concentration signals shown in Figure 5. In a multi-source environment, the odor pulses from one source will be intermixed with pulses from other sources. In such an environment, the average concentration of a compound is not a useful feature for olfactory scene analysis. Even if only one odor source is present, the statistical nature of the plume is such that several minutes of signal averaging are necessary to get an accurate estimate of average concentration. However, behavioral experiments in plumes of this sort indicate that animals make olfactory decisions on the order of a few seconds (Basil and Atema, 1994).

Like the visual and auditory systems, the olfactory system must be able to cope with wide ranges in signal (con-

Figure 5. Simulation of how an array of olfactory receptor cells might respond to the mingling of odor plumes from two different sources. The top two panels show the concentration signals from the two sources, and the bottom panel is the response from a simulation of 32 receptors that vary in their sensitivity to the two odorants.

that high molecular selectivity leads to poor temporal resolution—not unlike the trade-off between frequency selectivity and temporal resolution in auditory filters.

Olfactory receptors have been shown to respond rapidly enough that the temporal characteristics of the concentration signals could be available to the central nervous system (Gomez and Atema, 1996). Since most odors are mixtures and a single olfactory receptor cell can be stimulated by more than one compound, the odor from a single source will excite a number of different receptor cells, with the pattern of excitation varying from one odor mixture to another. In Figure 5 we simulate how an array of olfactory receptor cells might respond to the mingling of odor plumes from two different sources. The top two panels show the concentration signals from the two sources, and the bottom panel is the response from a simulation of 32 receptors that vary in their sensitivity to the two odorants. One can see from Figure 5 that, as in the auditory system, grouping can be done using temporal cues. In other words, receptors whose activities co-vary in time are likely to be responding to the same odor source.

Hardware models of olfactory scene analysis have not progressed very far due to the lack of sensors with the combination of appropriate chemical selectivity and fast temporal responses. Most current experiments are being done with surrogate odor sources for which fast sensors are available. The systems used in these experiments are generally designed to locate the odor source and not to classify the odor type. Due to the difficulty of accurately simulating chemical plumes in software, artificial systems for olfactory scene analysis often involve the use of robots. For example, we have used an aquatic robot (RoboLobster) that uses conductivity sensors to locate sources of salt in a freshwater flume (Grasso *et al.*, 2000).

Summary and Conclusions

The comparisons of strategies for scene analysis across the three sensory modalities—visual, auditory, and olfactory—described above illustrate several common themes that operate across modalities. For example, we see the dissection of the sensory signal, its processing in parallel pathways to extract key features, and then the grouping of portions of the signal to form perceptual objects. In all three of these senses, adaptation plays an important role in the first stages of processing. Fundamental trade-offs such as spectral *versus* temporal resolution or molecular selectivity *versus* temporal resolution shape peripheral processing. The mathematical concept of cross-correlation and its neural counterpart, coincidence detection, show up over and over again. We believe that by comparing strategies for sensory processing across sensory modalities as well as across many different species we can derive fundamental principles of sensory information processing that can be used to design artificial systems capable of analyzing complex environments.

Literature Cited

Basil, J., and J. Atema. 1994. Lobster orientation in turbulent odor plumes: simultaneous measurement of tracking behavior and temporal odor patterns. *Biol. Bull.* **187:** 272–273.

Boahen, K. A. 2000. Point-to-point connectivity between neuromorphic chips using address events. *IEEE Trans. Circuits Syst. II* **47:** 416–434.

Bregman, A. S. 1990. *Auditory Scene Analysis.* MIT Press, Cambridge, MA.

Colburn, H. S. 1996. Computational models of binaural processing. Pp. 332–400 in *Auditory Computation, Springer Handbook of Auditory Research,* H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay, eds. Springer Verlag, New York.

Echteler, S. M., R. R. Fay, and A. N. Popper. 1994. Structure of the mammalian cochlea. Pp. 134–171 in *Comparative Hearing. Mammals, Springer Handbook of Auditory Research,* R. R. Fay and A. N. Popper, eds. Springer Verlag, New York.

Fragniere, E., A. VanSchaik, and E. A. Vittoz. 1997. Design of an analogue VLSI model of an active cochlea. *Analog Integr. Circuits Signal Processing* **13:** 19–35.

Gomez, G., and J. Atema. 1996. Temporal resolution in olfaction: stimulus integration time of lobster chemoreceptor cells. *J. Neurophysiol.* **76:** 1340–1343.

Grasso, F. W. 2001. Invertebrate-inspired sensory-motor systems and autonomous, olfactory-guided exploration. *Biol. Bull.* **200:** 160–168.

Grasso, F. W., T. R. Consi, D. C. Mountain, and J. Atema. 2000. Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: progress and challenges. *Robot. Auton. Syst.* **30:** 115–131.

Hinck, T., and A. E. Hubbard. 1999. An analog VLSI image processing filter bank for enhancing oriented edge gradients. Pp. 897–902 in *Proc. 33rd Annual Conference on Information Sciences and Systems,* J. L. Prince and T. D. Tran, eds. Johns Hopkins University, Baltimore, MD.

Hoy, R. R. 1992. The evolution of hearing in insects as an adaptation to predation from bats. Pp. 115–139 in *The Evolutionary Biology of Hearing,* D. B. Webster, R. R. Fary, and A. N. Popper, eds. Springer Verlag, New York.

Hubel, D. H., and T. N. Wiesel. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160:** 106–154.

Javel, E. 1986. Basic response properties of auditory nerve fibers. Pp. 213–245 in *Neurobiology of Hearing,* R. Altschuler and R. Bobbin, eds. Raven Press, New York.

Lauffenburger, D. A., and J. J. Linderman. 1993. *Receptors—Models for Binding, Trafficking, and Signalling.* Oxford University Press, Oxford.

Lewis, B. 1992. The processing of auditory signals in the CNS of Orthoptera. Pp. 95–114 in *The Evolutionary Biology of Hearing,* D. B. Webster, R. R. Fary, and A. N. Popper, eds. Springer Verlag, New York.

Lyon, R., and S. Shamma. 1996. Auditory representations of timbre and pitch. Pp. 221–270 in *Auditory Computation, Springer Handbook of Auditory Research,* H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay, eds. Springer Verlag, New York.

Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman, San Francisco.

Mead, C. 1989. *Analog VLSI and Neural Systems.* Addison Wesley, Reading, MA.

Michelsen, A. 1992. Hearing and sound communication in small animals: evolutionary adaptations to the laws of physics. Pp. 61–77 in *The Evolutionary Biology of Hearing,* D. B. Webster, R. R. Fay, and A. N. Popper, eds. Springer Verlag, New York.

Michelsen, A. 1998. Biophysics of sound localization in insects. Pp. 18–62 in *Comparative Hearing: Insects, Springer Handbook of Auditory Research,* R. R. Hoy, A. N. Popper, and R. R. Fay, eds. Springer Verlag, New York.

Moore, P. A., and J. Atema. 1991. Spatial information in the three-dimensional fine structure of an aquatic odor plume. *Biol. Bull.* 181: 408–418.

Mountain, D. C., and A. E. Hubbard. 1996. Computational analysis of hair cell and auditory nerve processes. Pp. 121–156 in *Auditory Computation, Springer Handbook of Auditory Research,* H. L. Hawkins,

T. A. McMullen, A. N. Popper, and R. R. Fay, eds. Springer Verlag, New York.

O'Carroll, D. 1993. Feature-detecting neurons in dragonflies. *Nature* 362: 541–543.

Pasupathy, A., and C. E. Connor. 1999. Responses to contour features in macaque area V4. *J. Neurophysiol.* 82: 2490–2502.

Porat, M., and Y. Y. Zeevi. 1989. Localized texture processing in vision: analysis and synthesis in the Gaborian space. *IEEE Trans Biomed. Eng.* 36: 115–129.

Sarpeshkar, R., R. F. Lyon, and C. Mead. 1998. A low-power wide-dynamic-range analog VLSI cochlea. *Analog Integrated Circuits Signal Processing* 16: 245–274.

Webb, B., and T. Scutt. 2000. A simple latency-dependent spiking-neuron model of cricket phonotaxis. *Biol. Cybern.* 82: 247–269.

Weldon, T., and W. E. Higgins. 1999. Designing multiple Gabor filters for multitexture image segmentation. *Optical Eng.* 38: 1478–1489.

Zhang, S. W., M. V. Srinivasan, and T. Collett. 1995. Convergent processing in honeybee vision: multiple channels for the recognition of shape. *Proc. Notl. Acad. Sci. USA* 92: 3029–3031.